An Analysis of Approval-Based Committee Rules for 2D-Euclidean Elections

Michał Godziszewski,[1] Paweł Batko,[2] Piotr Skowron,[1] Piotr Faliszewski[2]
[1] University of Warsaw, Poland, [2] AGH University, Kraków, Poland
mtgodziszewski@uw.edu.pl, batkop@agh.edu.pl, p.skowron@mimuw.edu.pl, faliszew@agh.edu.pl

Abstract

We study approval-based committee elections for the case where the voters' preferences come from a 2D-Euclidean model. We consider two main issues: First, we ask for the complexity of computing election results. Second, we evaluate election outcomes experimentally, following the visualization technique of Elkind et al. (2017). Regarding the first issue, we find that many NP-hard rules remain intractable for 2D-Euclidean elections. For the second one, we observe that the behavior and nature of many rules strongly depend on the exact protocol for choosing the approved candidates.

1 Introduction

The idea of committee elections is that a group of agents (typically referred to as the voters) wants to choose a fixed-sized subset of candidates (typically referred to as a committee). For example, the voters may be choosing the finalists of some competition, the members of some governing body, or the products to offer in an online store (see, e.g., the overview of Faliszewski et al. (2017) for a discussion of various types of committee elections). The committee should reflect the preferences of the voters, and its selection should follow the principles underlying its purpose (for example, the finalists of a competition should be individually excellent, members of a governing body should represent the voters proportionally, and a store's portfolio should be diverse).

We consider the approval preference model, i.e., we assume that each voter specifies a subset of candidates that he or she finds suitable for the committee, and we focus on the case where these approval sets are derived from some (two dimensional) Euclidean model. We do so for two reasons. First, the results of 2D-Euclidean elections can be easily interpreted and we want to verify if various approval-based committee rules (ABC rules, for short) indeed implement the desired principles in choosing the committees. To this end, we use the visualization technique of Elkind et al. (2017). Second, for many prominent ABC rules it is known that computing their results is NP-hard in general, but becomes tractable if the preferences are, in some sense, one-dimensional (for example, all the domain restrictions considered by Elkind and Lackner (2015) are one-dimensional). We check if these polynomial-time results can be extended to the two-dimensional case.

We consider voting rules that seek committees of different types. In particular, we consider Multiwinner Approval Voting (AV), which focuses on individual excellence, Proportional Approval Voting (PAV), Phragmén's Sequential rule (Phr), and Rule X, which focus on proportionality, and Approval Chamberlin–Courant (CC), which focuses on diversity. Additionally, we also consider Minimax Approval Voting (MAV), which is based on the egalitarian principle.

Approval-Based Euclidean Elections.
Briefly put, in a Euclidean model each candidate and each voter is represented as his or her *ideal point*, i.e., a point in some Euclidean space $\mathbb{R}^\ell$, whose coordinates are interpreted as a given candidate's or voter's positions on some ℓ issues (for example, in a two-dimensional model these two issues may correspond to the extents to which an individual supports personal and economic freedom). To derive the voters' approval sets, we use the following two principles:

1. In the voter-range model, if a voter approves some candidate then he or she also approves all

the closer ones.

2. In the candidate-range model, if a candidate is approved by a voter, all the voters closer to this candidate also approve him or her.

For the one-dimensional case, these models correspond to the candidate interval and voter interval models of Elkind and Lackner (2015). We also consider what we call a *voter/candidate range* model, which generalizes both approaches, and using it we generalize/strengthen some results from the literature.

Main Contributions.
We seek to understand the structure of the committees produced by the considered rules and whether these committees can be computed efficiently.

In our first set of results, we show that all our rules that are NP-hard in the general approval setting (e.g., MAV, CC, and PAV) remain NP-hard for 2D-Eucludiean elections, both for the voter and candidate range models. Our proofs hold under the assumption that we are given the ideal points and radii of the candidates and voters; this is important as otherwise even recognizing if an election comes from a 2D-Euclidean domain may be hard (indeed, Peters (2017) has shown it for the ordinal case; we expect the same for the approval one).

In the second set of results, we present visualizations of our ABC rules under several distributions of the ideal points and several strategies for choosing the voters' or candidates' radii (we employ the technique of Elkind et al. (2017), which shows how frequently committee members are selected from given areas of the preference space). We obtain several high-level conclusions. For example, we find that forcing the voters to approve a given number of candidates leads to unappealing results, or that AV might have some deficiencies as a rule to choose individually excellent candidates.

2 Preliminaries

For an integer p, by $[p]$ we mean the set $\{1, \ldots, p\}$. We use the Iverson bracket notation, i.e., for a logical formula F, by $[F]$ we mean 1 if F is true and 0 otherwise.

Approval Elections.
An *approval-based election* (in short, an *election*) is a pair $E = (C, V)$, where $C = \{c_1, \ldots, c_m\}$ and $V = \{v_1, \ldots, v_n\}$ are, respectively, the set of candidates and the set of voters. For each voter $v \in V$, by $A(v)$ we denote the *approval set* of v, i.e., the set of those candidates that voter v finds acceptable. Conversely, by $V(c)$ we denote the set of voters who approve candidate c, i.e., $V(c) = \{v \in V \mid c \in A(v)\}$. To make the notation lighter, we assume that the approval sets of the voters are implicitly included in the elections.

Elkind and Lackner (2015) introduced a number of domain restrictions regarding the voters' preferences. For example, if it is possible to order the candidates so that each voter approves their contiguous subset, then we say that the voters have *candidate interval (CI)* preferences. Similarly, if it is possible to order the voters so that each candidate is approved by a contiguous group of voters, then we speak of *voter interval (VI)* preferences. We focus on Euclidean preferences, which we discuss in Section 3.

ABC Rules.
An *approval-based committee rule*—in short an *ABC rule*—is a function that given an election (C, V) and a positive integer $k \in \mathbb{N}$, returns a nonempty family of size-k subsets of C, referred to as the winning committees. For a general overview of ABC rules, we point the reader to the recent survey by Lackner and Skowron (2020). Below we describe the rules that we focus on (we let $E = (C, V)$ be an election and k be the desired committee size).

Thiele Methods. Fix a non-decreasing function $w \colon \mathbb{N} \to \mathbb{R}$. The w-score of a committee W is defined as:

$$w\text{-score}(W) = \sum_{v \in V} w(|W \cap A(v)|).$$

The w-Thiele rule returns the committees with the maximal w-scores. Rules of this type (often referred to as Thiele methods) were introduced by Thiele (1895). Notable examples of Thiele rules include Multiwinner Approval Voting (AV), Approval Chamberlin-Courant rule (CC), and Proportional Approval Voting (PAV), defined, respectively, through the following w-functions:

$$w_{\mathrm{AV}}(t) = t; \quad w_{\mathrm{CC}}(t) = [t \geq 1]; \quad w_{\mathrm{PAV}}(t) = \sum_{i=1}^{t} 1/i.$$

AV is focused on individual excellence, CC gives diverse committees, and PAV seeks proportional representation.

Phragmén's Sequential Rule. The rule starts with an empty committee and extends it until k candidates are found: Each candidate costs one dollar, the voters earn virtual money at some fixed rate, e.g., one dollar per second (the time is continuous), and as soon as some voters can buy a not-yet-selected candidate c that they all approve, the rule includes c in the committee and resets their budgets to 0. This rule was introduced by Phragmén (1894) and seeks proportional representation of the voters.

Rule X. This is a two-phase rule, where both phases resemble the Phragmén's Sequential Rule, but in the first one the voters get their money upfront, and in the second one they earn it as in Phragmén's Sequential Rule, but their starting budgets depend on the first phase. For a formal definition of the rule, we refer to the paper of Peters and Skowron (2020), where the rule is introduced.[1] In this paper we only consider the first phase of the rule, which can return committees of size strictly smaller than k.

Minimax Approval Voting (MAV). Given two subsets of candidates, $A, B \subseteq C$, their Hamming distance is $d_{\mathrm{Ham}}(A, B) = |A \setminus B| + |B \setminus A|$. MAV, introduced by Brams et al. (2007), selects committees that minimize the Hamming distance to the farthest vote, i.e., the committees W that minimize $\max_{v \in V} d_{\mathrm{Ham}}(A(v), W)$. Thus the rule implements the egalitarian principle; for other such rules, see, e.g., the works of Betzler et al. (2013) and Aziz et al. (2018).

Complexity of Winner Determination.
The outcomes of AV, Phragmén's Sequential Rule, and Rule X are computable in polynomial time, provided that we break ties according to some simple rule (e.g., lexicographically). For MAV and all w-Thiele rules with non-linear, concave w functions it is NP-hard to decide if there is a committee that achieves at least a given score (see the works of LeGrand (2004), Procaccia et al. (2008), and Aziz et al. (2015) for the cases of MAV, CC, and PAV, respectively, and the work of Skowron et al. (2016) for a general result regarding Thiele rules). Yet, for MAV there are polynomial-time algorithms for the candidate and voter interval cases Liu and Guo (2016), and for Thiele rules defined by concave w-functions there are polynomial-time algorithms for the candidate interval case Peters and Lackner (2020); the problem is open for the other Thiele rules and for the voter interval case.

Our rules have also been studied with respect to approximation and parameterized complexity; we point the reader to the survey of Lackner and Skowron (2020) for these results.

[1] We omit the definition due to restricted space. To appreciate our results, it suffices to know that Rule X is similar to Phragmén's Sequential Rule but provides stronger proportionality guarantees.

3 Euclidean Preferences

In this section we describe our model of Euclidean-based approval preferences. While such models have been studied for over half a century (see, e.g., the works of Davis and Hinich (1966), Plott (1967), Enelow and Hinich (1984, 1990) for some early discussions), researchers mostly focused on ordinal preferences, and when they considered the approval setting, they usually analyzed probabilistic models (see, e.g., the work of Laslier (2006)).

Given an election $E = (C, V)$, we say that the voters have tD-Euclidean preferences ($t \in \mathbb{N}$) if for each agent $a \in C \cup V$ (i.e., for each candidate and each voter) there exists a point $x_a = (x_{a,1}, \dots, x_{a,t})$ in $\mathbb{R}^t$ and a nonnegative real value $r_a \in \mathbb{R}$ such that:

$$c \in A(v) \iff \sqrt{\sum_{j=1}^{t}(x_{c,j} - x_{v,j})^2} \le r_c + r_v.$$

Intuitively, for $a \in C \cup V$ the point x_a describes a's ideal position in a t-dimensional space of opinions. For a candidate $c \in C$, r_c can be seen as c's charisma: It specifies which positions surrounding his or her ideal one the candidate can accommodate credibly. For a voter $v \in V$, r_v specifies v's willingness to compromise, i.e., the positions around his or her ideal one that the voter still accepts. Two special cases of Euclidean Preferences are:

1. The voter range model (VR), where we require that all the candidates have radii equal to zero.
2. The candidate range model (CR), where all the voters have radii equal to zero.

We refer to the full model as the voter/candidate range model (VCR). Elkind and Lackner (2015) argued that the candidate interval model is equivalent to our 1D-VR model (although they used a different name, of course). It turns out that the voter interval model is equivalent to our 1D-CR model.[2]

Proposition 1. *The sets of voter interval and 1D-CR elections are equal.*

The VCR model is strictly more powerful than the VR and CR ones. For example, election with candidate set $C = \{a, b, c, d\}$ and voters with approval sets $\{a, b\}$, $\{b, c\}$, $\{b, d\}$, and $\{a, b, c, d\}$ is VCR, but neither VR nor CR.

4 Computing Winning Committees

The main goal in this section is to show that our NP-hard rules remain intractable even in the 2D-Euclidean setting. Yet, first we briefly consider 1D-Euclidean elections.

4.1 One-Dimensional Euclidean Preferences

Elkind and Lackner (2015) have shown polynomial-time algorithms for computing CC winning committees in 1D-VR and 1D-CR elections. We unify this result into a single algorithm for 1D-VCR elections (the main idea is remarkably close to that for the 1D-VR case).

Proposition 2. *There is an algorithm that given a 1D-VCR election and committee size, computes some winning CC committee in polynomial time.*

A natural question is whether it is possible to extend the above result to PAV, other Thiele methods. We leave this question open. It is also interesting to consider MAV as in this case 1D-VR and 1D-CR algorithms do exist Liu and Guo (2016).

[2]It is a bit confusing that the voter range corresponds to the candidate interval model and candidate range corresponds to voter interval. This crossing is due to the fact that our terminology regards the reason for approval, and the terminology of Elkind and Lackner (2015) regards the shape of approval sets.

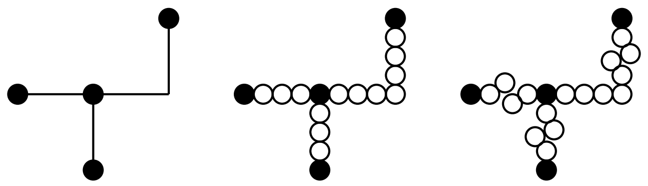


Figure 1: Construction of a penny graph.

4.2 Two-Dimensional Euclidean Preferences

In our NP-hardness proofs for the 2D Euclidean elections we use *penny graphs*. A penny graph is defined by a set of unit disks, i.e., balls of diameter one in $\mathbb{R}^2$, such that no two disks overlap (but they can touch). Each disk corresponds to a vertex and two vertices are connected by an edge if their disks touch (i.e., if their centers are exactly at distance 1). A graph is a penny graph if it has such a representation by unit disks (the name comes from the analogy between the disks and pennies lying on a flat surface). All penny graphs are planar. We will need the following algorithm of Valiant.

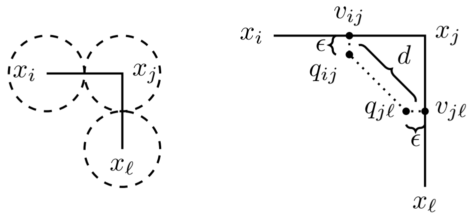
Lemma 1. *Valiant (1981). There is a polynomial-time algorithm that given a planar graph with maximum degree at most 4 computes its embedding on the plane so that its vertices are at integer coordinates and its edges are drawn with vertical and horizontal line segments.*

Recall that in the Independent Set problem (IS) we are given a graph $G = (X, E)$ and a positive integer r. We ask if there exists an independent set of G—i.e., a subset of vertices $U \subseteq X$ such that no two vertices in U are adjacent—of size at least r. It is known that the problem is NP-hard for cubic planar graphs (Mohar, 2001, Theorem 4.1(a)). Given an instance (G, r) of IS, where G is a cubic planar graph, we construct an instance of IS for penny graphs, as follows (we use the construction of Cerioli et al. (2011, Theorem 1.2); we repeat it here as we need its specific properties).

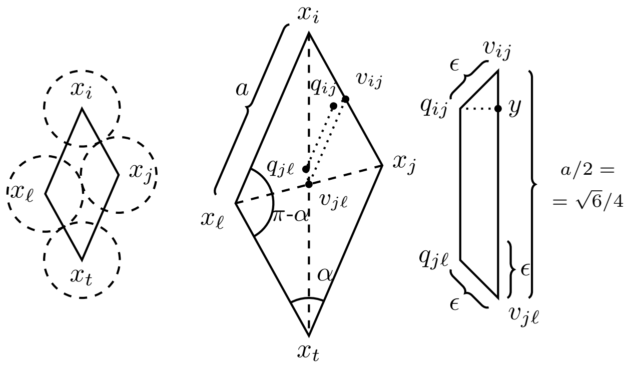
First, we use Lemma 1 to obtain a planar representation of G, where the vertices are at integer coordinates and the edges consist of vertical and horizontal line segments (see the left-hand side of Figure 1; note that in this figure the vertices have degrees at most three, not necessarily exactly three). Second, we multiply vertex coordinates by four, ensuring that the lengths of the line segments forming the edges also are multiples of four. Third, for each vertex v we put a unit disk centered at the position of v, and we replace all the line segments forming the edges by sequences of consecutive unit disks (located on the integral points within these lines; see the center of Figure 1). This way, each edge becomes a sequence of $4t - 1$ disks, where t is an integer (possibly different for each edge). Finally, for each edge we introduce a single *local displacement*, which consists of replacing the second disc that lies on the edge with two tangent disks (it does not matter from which end we start counting the disks); these two disks are also tangent to the disks on the two sides of the disk that we replaced (see the right-hand side of Figure 1). Local displacements ensure that disks on the edges come in multiples of four. All in all, we obtain a penny graph.

Let G' be the penny graph that we constructed. Each vertex of G' has either two or three adjacent vertices. The vertices with two neighbors correspond to disks put on the edges and we refer to them as *intermediate*. We call a vertex locally displaced if it corresponds to a disk that was introduced as a result of a local displacement. Let L be the total number of intermediate vertices. One can easily verify that G has an independent set of size r if and only if G' has an independent set of size $r' = r + L/2$ (this follows from the work of Cerioli et al. (2011)). We refer to the penny graphs obtained by this construction as *almost integral* and we use the fact that IS is NP-hard for them.

We are ready to show that for a large class of Thiele rules, computing the results is intractable even for 2D elections.



(a) Case 1: $\overline{x_i x_j}$ and $\overline{x_j x_\ell}$ are orthogonal.



(b) Case 2: Disks at x_j and x_ℓ form a local displacement.

Figure 2: Two cases for estimating the distance between points q_{ij} and $q_{j\ell}$ in the proof of Theorem 1.

Theorem 1. *For each non-linear concave function $w\colon \mathbb{N} \to \mathbb{R}$, deciding if there is a committee of a given size with at least a given w-Thiele score is* NP*-hard for 2D-VR elections, even if the voters have the same approval radii.*

Proof. Let p be the largest integer such that for each $p' \in [p]$ we have that $w(p') = p' \cdot w(1)$. Since w is non-linear, p is well-defined. Since w is concave, $w(p+1) < (p+1) \cdot w(1)$. Further, we fix ϵ to be a small positive constant—the upper-bound on the value of ϵ will be clear from the construction.

We reduce from the Independent Set problem for almost integral penny graphs (where the graph is given by its geometric representation). Let (G, r) be an instance of IS, where G is an almost integral penny graph and r is an integer (these are G' and r' from the construction above). Let n denote the number of edges in G. We distinguish the following points in $\mathbb{R}^2$ (we will use them as the ideal points of the agents):

Vertex Points: For each vertex x_i, we have a vertex point located in the center of x_i's disk; we overload the notation and also refer to this point as x_i.

Edge Points: For each edge $e = \{x_i, x_j\}$ in G, we have a point in the middle of e, to which we refer as e_{ij} (we view edges in G as straight, unit-length line segments).

Bisector Points: For each edge $\{x_i, x_j\}$ in G, we take the bisector of the line segment $\overline{x_i x_j}$ and let q_{ij} be a point on its bisector (on an arbitrary side) such that its distance from the line segment is ϵ.

Given (G, r), we construct an election $E = (C, V)$ with the following candidates and voters:

1. The set of candidates consists of *vertex candidates* and *bisector candidates*: In each vertex point x_i we put one vertex candidate, called c_i, and in each bisector point we put $p-1$ bisector candidates, called $b_{ij}^1, \ldots, b_{ij}^{p-1}$. We write C_b to denote the set of all bisector candidates.

2. We have the following three groups of voters (we argue that the approval sets are specified correctly a bit later):

   (a) The edge voters: For each edge $e = \{x_i, x_j\}$ we have a voter located in point e_{ij}, who approves c_i, c_j, and $p-1$ bisector candidates located in q_{ij}.

   (b) The vertex voters: For each intermediate vertex x_i, we have one voter, v_i, who is located in point x_i and approves only c_i.

   (c) The bisector voters: In each bisector point q_{ij} we put three voters, $u^1_{ij}, u^2_{ij}, u^3_{ij}$, who all approve the $p-1$ bisector candidates that are in their location.

We set the committee size to be $k = r + (p-1)n$. We ask if there exists a committee with w-score greater or equal to $s = (3r + 4n(p-1))\,w(1)$. This ends the construction.

Let us now show that if all the voters have approval radius $1/2$ then they approve exactly the candidates indicated above. First, observe that for each vertex x_i of the penny graph, the distance between point x_i and each bisector point of the form q_{ij} (i.e., each bisector point associated with an edge incident to x_i) is $\sqrt{1/4 + \epsilon^2} > 1/2$. Thus the distances between the vertex voters and the bisector candidates, as well as between the vertex candidates and the bisector voters, are strictly greater than $1/2$, and the respective approval ballots do not interfere with each other. Further, it is clear that if ϵ is sufficiently small, then for each q_{ij}, its distance from each x_ℓ is also strictly greater than $1/2$.

It remains to consider balls of diameter 1, centered at some bisector points, q_{ij} and $q_{j\ell}$. There are two cases to analyze. The first one occurs when the line segments $\overline{x_i x_j}$ and $\overline{x_j x_\ell}$ are orthogonal (see Figure 2a). In such a case the distance between q_{ij} and $q_{j\ell}$ is $(1/2 - \epsilon)\sqrt{2}$, which is more than $1/2$ for $\epsilon < 1/2 \cdot (1 - \frac{\sqrt{2}}{2}) \approx 0.14$ (so we require $\epsilon < 0.14$).

The second case occurs when the disks centered at x_j and x_ℓ constitute a local displacement (see Figure 2b, two plots on the left). Then the points x_i, x_j, x_ℓ, and x_t form a parallelogram, where the lengths of sides $\overline{x_i x_j}$ and $\overline{x_\ell x_t}$ are 1, and the lengths of sides $\overline{x_i x_\ell}$ and $\overline{x_j x_t}$ are $a = 1/2 \cdot \sqrt{6}$. To see why this is the case, note that the lengths of the diagonals of the parallelogram are 1 (the diagonal $\overline{x_j x_\ell}$) and 2 (the diagonal $\overline{x_i x_t}$), and let α be the magnitude of the angle $\angle x_j x_i x_\ell$. By the law of cosines we have. $1^2 + a^2 - 2a\cos(\alpha) = 1^2$, and $1^2 + a^2 - 2a\cos(\pi - \alpha) = 2^2$. After adding these two equalities and simple calculations we obtain that $a = 1/2 \cdot \sqrt{6}$. Now observe that the points $v_{ij}, v_{j\ell}$, q_{ij}, and $q_{j\ell}$ form an isosceles trapezoid (see the right plot in Figure 2b). If $\epsilon < 1/2 \cdot \left(\frac{\sqrt{6}}{4} - \frac{1}{2}\right)$, then clearly the distance between q_{ij} and $q_{j\ell}$ is greater than $1/2$. This completes the proof that the voters' approval sets are indeed as indicated in the construction.

It remains to show that the reduction is correct. Let us assume that G is a "yes" instance of IS. Take any independent set U of size r and define the corresponding committee to be $S_U = \{c_i \mid x_i \in U\} \cup C_b$, i.e., let S_U consist of the vertex candidates corresponding to the members of U and all the bisector candidates. Let us calculate the w-score of S_U:

1. Each bisector candidate b^ℓ_{ij} contributes exactly $4w(1)$ points to the committee. This is because he or she is approved by three bisector voters u^1_{ij}, u^2_{ij}, and u^3_{ij} at her location (and each of these voters approves exactly $p-1$ bisector candidates), and by the edge voter at point e_{ij} (who in total approves at most p committee members; the $p-1$ bisector candidates and, as U is an independent set, at most one vertex candidate). Each of these voters contribues $w(1)$ points for each bisector candidate.

2. Each vertex candidate from the committee contributes $3w(1)$ points. Indeed, each of them is approved by three voters (either three edge voters or two edge voters and one vertex voter). Each vertex voter approves exactly one candidate and, by the argument from the previous point, each edge voter approves at most p committee members. Thus each vertex candidate from the committee brings in $w(1)$ points for each voter that approves him or her.

(a) The uniform square model (b) The asymmetric Gaussians model (c) The overlapping squares model

Figure 3: Density functions for our models of generating ideal points (red areas correspond to candidates, green areas correspond to voters, and olive areas correspond to both).

So the w-score of S_U is $(3r + 4n(p-1))w(1)$, as required.

For the other direction, assume that there is a size-k committee S with score at least $s = (3r + 4n(p-1))w(1)$. Since each vertex candidate is approved by exactly three voters and, thus, can contribute at most $3w(1)$ points to the score of the committee, if S has score s then it must include all the bisector candidates and each of these bisector candidates has to contribute $4w(1)$ points to the committee. The latter happens exactly if all the edge voters approve at most p committee members, which happens exactly if the vertex candidates from S form an independent set. $\square$

The above proof also holds for the 2D-CR model (it suffices to assume that the candidates have radii $1/2$ and the voters have radii 0). An analogous result also holds for MAV.

Theorem 2. *Deciding if there is a MAV committee of a given size and with at most a given score is* NP*-hard for 2D-VR elections (2D-CR elections) even if all voters (all candidates) have the same approval radius and each voter approves the same number of candidates.*

Theorems 1 and 2 also hold for higher dimensions.

5 Visualization

The visualization idea of Elkind et al. (2017) is to generate a large number of elections, where the agents' ideal points come from some distributions, compute their results, and draw 2D histograms indicating how many winners appear in each area of the space. We adopt their methodology to the approval-based setting.

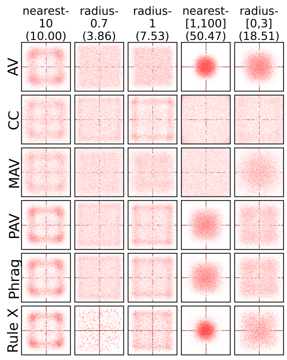
5.1 Generating the Histograms

We consider three ways to generate the points representing the candidates and voters (illustrated in Figure 3):

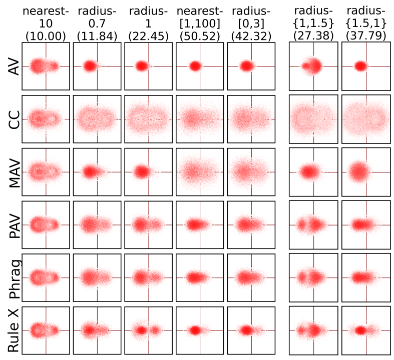
1. The *uniform square* model, where the points are selected uniformly at random from the $[-3, 3] \times [-3, 3]$ square. This is a "baseline" distribution that was also considered by Elkind et al. (2017).

2. The *asymmetric Gaussians* model, where 70% of the points are generated from a two-dimensional Gaussian distribution with center $(-1, 0)$ and standard deviation 0.8, and 30% come from a Gaussian with center $(1, 0)$ and the same standard deviation. This model simulates a society where a large majority has views centered in one area and a significant minority has views centered in some distance from them. This model is intended to highlight rules' abilities to choose proportional results.

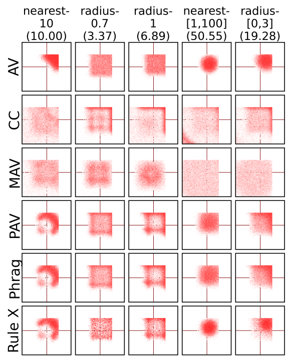
3. The *overlapping squares* model, where the points of the voters are selected uniformly at random from square $[-1.5, 3] \times [-1.5, 3]$ and the points of the candidates are selected uniformly



(a) The uniform square model



(b) The asymmetric Gaussians model



(c) The overlapping squares model

Figure 4: Histograms for our rules and ways of generating elections. The numbers in parentheses over each column provide the average number of candidates approved by a single voter.

at random from square $[-3, 1.5] \times [-3, 1.5]$. This model captures a setting where the populations of the candidates and voters represent different opinions (it is quite extreme in this respect, which makes the differences between the rules more visible).

In most settings we focus on the voter range model, where the candidates have radii set to zero and voters' radii are generated using models from one of the following two groups:

1. In the first group, we either fix the number of approved candidates or the approval radius. Specifically, in the 10-*nearest* model, each voter's radius is such that he or she approves the 10 closest candidates. In the *radius*-0.7 and *radius*-1 models, the radii are fixed to, respectively, 0.7 and 1. We refer to these three models as the *fixed* ones.

2. In the second group, we choose the number of approved candidates or the approval radius from a uniform distribution. Specifically, in the $[1, 100]$-*nearest* model, for each voter we choose number t uniformly at random from the set $\{1, \dots, 100\}$ and then select the radius so that the voter approves exactly t closest candidates. In the *radius*-$[0, 3]$ model, for each voter we choose the radius uniformly at random from interval $[0, 3]$. We refer to these two models as the *uniform* ones.

For the asymmetric Gaussians model, we also consider the candidate range model, with the following candidate radii:

3. In the *radius*-$\{1, 1.5\}$ model, the candidates from the left-hand side Gaussian (the larger one) have radius 1 and the other ones have radius 1.5. In the *radius*-$\{1.5, 1\}$ model, these values are swapped.

To draw a histogram for a given voting rule and models of generating agents' points and radii, we proceed as follows. First, we generate 2000 elections with 100 candidates and 100 voters each. Then we compute their winning committees of size 10.[3] Next, we consider the $[-3, 3] \times [-3, 3]$ square partitioned into cells of size 0.05×0.05 and for each cell we count how many members of the winning committees fall there. Finally, we plot the thus-obtained 2D histograms. We map the numbers of committee members in each cell to color intensities using the formula of Elkind et al. (2017); the darker a cell, the more committee members it contains.

[3] We used implementations from the `abcvoting` library (https://github.com/martinlackner/abcvoting) in *irresolute* mode (except Rule X) and broke ties uniformly at random. This library uses integer linear programming formulations of the respective NP-hard rules; by Theorems 1 and 2 we know that this approach is as reasonable in the 2D Euclidean setting as in the general one.

5.2 Discussion

We present the histograms for our voting rules, models of generating agent's points, and models of generating the radii in Figure 4. Below we analyze the obtained results.

Choosing Radii Matters.
The most obvious—but, perhaps, the most important—observation from Figure 4 is that depending how we choose the voters' radii, the results can vary greatly for all our rules. There are several reasons for this. Foremost, it is natural that the results for the fixed models are different than those for the uniform ones. For example, Bredereck et al. (2020, Lemma 9) show how to use Thiele rules to simulate corresponding committee scoring rules Skowron et al. (2019); Faliszewski et al. (2019) with Borda as the underlying scoring function. The nearest-$[1, 100]$ radius model is a randomized variant of their construction. As a consequence, in this model AV behaves like the classic k-Borda rule and PAV behaves like the HarmonicBorda rule (indeed, for these cases their histograms are identical to those obtained by Elkind et al. (2017) for k-Borda and HarmonicBorda). On the other hand, for the nearest-10 radii model, AV is known as the Bloc rule and is known to behave very differently from k-Borda.

Fixed Models.
It is intriguing to compare the histograms for the nearest-10 model and the radius-0.7 and radius-1 models. In the former, each voter has to approve 10 closest candidates, even if they are quite far away from him or her, whereas in the two latter ones, the voters only approve candidates that are close (we chose radii 0.7 and 1 because in the uniform square model the latter leads to approving close to 8 candidates on average, and in the asymmetric Gaussians model the former leads to approving under 12 candidates on average; we view these values as close enough to 10 so that the models are comparable to nearest-10, and as sufficiently different from each other to be interesting). In most cases, the histograms for the nearest-10 model have very pronounced artifacts, which seem to go against the spirit of the respective rules. E.g., in the uniform square model all rules have darker areas in the corners, even though there is no reason to consider the candidates there as more appealing than the other ones. Similarly, in the asymmetric Gaussians model, PAV and Phragmén choose fewer candidates from the center of the larger Gaussian, even though many agents have ideal points in this area, and one would expect a proportional rule to choose *more* candidates from there. These artifacts either disappear or are less pronounced in the radius-0.7 and radius-1 models. This suggests that if one is using approval-based voting rules, then there should be no fixed number of candidates that the voters should approve.

Elkind et al. (2017) also observed such artifacts for the Bloc rule, which is equivalent to the AV rule in the nearest-10 model (or, more specifically, in the nearest-k model, where k is the committee size). Our results suggest that these artifacts appear due to requiring the voters to approve candidates that are located far away from them (which happens when the number of to-be-approved candidates is fixed).

Uniform Models.
The differences between the $[1, 100]$-nearests and radius-$[0, 3]$ models are less worrisome than those between the fixed models, even if sometimes quite visible; see, e.g., PAV and Phragmén in the overlapping squares model. Yet, the results for radius-$[0, 3]$ are more appealing as more candidates in the top-right corner of the candidate square are selected, closer to a large group of voters.

PAV Versus Phragmén.
One striking observation is that the histograms for PAV and Phragmén are, in essence, indistinguishable for all our settings. One could argue that this is natural because both rules aim at achieving proportional representation. However, axiomatic studies suggest that they understand proportionality in quite different ways Peters and Skowron (2020). Our histograms suggest that in the 2D Euclidean models these two ways coincide.

CC, MAV, and Diversity.

Generally, the histograms for CC show fairly uniform coverage of the candidates' and voters' views. This is good behavior for a rule that aims at choosing diverse committees. The MAV rule often behaves quite similarly (especially in the uniform square model) but can also follow AV's behavior. The similarity to CC—and, more generally, tendency to select diverse committees—can be explained by its egalitarian nature: If all the voters approve a similar number of candidates, then MAV seeks a committee that maximizes the number of its members approved by the worst-off voter (i.e., the voter who approves fewest of them). The similarity to AV appears in settings where there is a large group of voters who approve many more candidates than the remaining voters (as in the asymmetric Gaussians with radius-0.7 or 1): If a voter approves many candidates, then MAV has to put many of them in the committee, and if there are many voters like this (approving many common candidates), the result is similar to using AV.

AV And Individual Excellence.
AV is typically seen as a rule for choosing individually excellent candidates. Our histograms indicate that, at least in some settings, it might not do well in this task (or, in a different interpretation, approval sets might be insufficient for it). To this end, let us focus on the uniform square model. Faliszewski et al. (2017) argue that in the individual excellence setting, similar candidates should be treated similarly (two similar candidates should either both be in the committee or both be out, up to boundary cases). If we take geometric proximity of candidates' ideal points as a measure of similarity, then an individually excellent committee should consist of candidates located in the center (by the given argument, the selected candidates should be close to each other; center location follows from symmetry). Yet, AV achieves such histograms only for nearest-$[1, 100]$ and radius-$[0, 3]$ models. Since these models, effectively, act as if the voters ranked the candidates,[4] it is an argument that in some settings ordinal models are better suited for individual excellence than the approval ones.

Rule X And Proportionality.
We note that the histograms for Rule X (or, rather, its first part; recall Section 2) are between those for Phragmén and AV. (see, e.g., the asymmetric Gaussians case). This is surprising as Peters and Skowron (2020) have shown that even the first part of Rule X alone has strong proportionality guarantees Peters and Skowron (2020) and one would not expect increased similarity to AV.

Candidate Range Models.
Finally, let us consider the asymmetric Gaussian model and candidate range models, i.e., radius-$\{1, 1.5\}$ and radius-$\{1.5, 1\}$ models. As expected for a diversity-oriented rule, CC behaves fairly similarly in both cases. PAV and Phragmén, on the other hand, are quite asymmetric. For the former model there is a visible area between the centers of the Gaussians from which no candidates are selected. For the latter model, no such phenomenon occurs. One explanation is that in the radius-$\{1, 1.5\}$ model, the more charismatic candidates from the smaller Gaussian are too few to satisfy the voters from the left-hand extreme of the larger Gaussian, and these voters are sufficiently numerous to elect the less charismatic, but closer, candidates (to some extent, this is supported by the shape of the histogram for AV in this model).

6 Conclusions

We have shown that many NP-hard approval-based committee rules remain NP-hard in the 2D Euclidean setting, even if in 1D settings they can by computed in polynomial time. We have also computed visualizations of our rules and made multiple observations, the crucial one being that one should not force the voters to approve a fixed number of candidates.

[4]Consider, e.g., the nearest-$[1, 100]$ model and a group of voters located close to each other. The candidates approved by all of them are the closest to the group, and, generally, the fewer voters approve a given candidate, the farther he or she is. This way AV gets similar information as if the voters ranked the candidates.

Acknowledgments

M. Godziszewski was supported by the Foundation for Polish Science within the Homing programme (Project title: "Normative Comparison of Multiwinner Election Rules") P. Batko and P. Faliszewski were supported by the funds assigned to the AGH University by the Polish Ministry of Science and Higher Education. P. Skowron was supported by the National Science Centre (project 2019/35/B/ST6/02215).

References

Aziz, H., Faliszewski, P., Grofman, B., Slinko, A., and Talmon, N. (2018). Egalitarian committee scoring rules. In *Proceedings of IJCAI-2018*, pages 56–62.

Aziz, H., Gaspers, S., Gudmundsson, J., Mackenzie, S., Mattei, N., and Walsh, T. (2015). Computational aspects of multi-winner approval voting. In *Proceedings of AAMAS-2015*, pages 107 – 115.

Betzler, N., Slinko, A., and Uhlmann, J. (2013). On the computation of fully proportional representation. *Journal of Artificial Intelligence Research*, 47:475–519.

Brams, S. J., Kilgour, D. M., and Sanver, M. R. (2007). A minimax procedure for electing committees. *Public Choice*, 132(3–4):401–420.

Bredereck, R., Faliszewski, P., Kaczmarczyk, A., Knop, D., and Niedermeier, R. (2020). Parameterized algorithms for finding a collective set of items. In *Proceedings of AAAI-2020*, pages 1838–1845.

Cerioli, M., Faria, L., Ferreira, T., and Protti, F. (2011). A note on maximum independent sets and minimum clique partitions in unit disk graphs and penny graphs: complexity and approximation. *RAIRO - Theoretical Informatics and Applications - Informatique Théorique et Applications*, 45(3):331–346.

Davis, O. and Hinich, M. (1966). A mathematical model of preference formation in a democratic society. In Bernd, J. L., editor, *Mathematical Applications in Political Science II*, pages 175–208. Southern Methodist University Press.

Elkind, E., Faliszewski, P., Laslier, J., Skowron, P., Slinko, A., and Talmon, N. (2017). What do multiwinner voting rules do? An experiment over the two-dimensional euclidean domain. In *Proceedings of AAAI-2017*, pages 494–501.

Elkind, E. and Lackner, M. (2015). Structure in dichotomous preferences. In *Proceedings of IJCAI-2015*, pages 2019–2025.

Enelow, J. and Hinich, M. (1984). *The spatial theory of voting: An introduction*. Cambridge University Press.

Enelow, J. and Hinich, M. (1990). *Advances in the spatial theory of voting*. Cambridge University Press.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2017). Multiwinner voting: A new challenge for social choice theory. In Endriss, U., editor, *Trends in Computational Social Choice*. AI Access.

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2019). Committee scoring rules: Axiomatic characterization and hierarchy. *ACM Transactions on Economics and Computation*, 6(1):Article 3.

Lackner, M. and Skowron, P. (2020). Approval-based committee voting: Axioms, algorithms, and applications. Technical Report arXiv:2007.01795 [cs.GT], arXiv.org.

Laslier, J. F. (2006). Spatial approval voting. *Political Analysis*, 14(2):160–185.

LeGrand, R. (2004). Analysis of the minimax procedure. Technical Report WUCSE-2004-67, Department of Computer Science and Engineering, Washington University.

Liu, H. and Guo, J. (2016). Parameterized complexity of winner determination in minimax committee elections. In *Proceedings of AAMAS-2016*, pages 341–349.

Mohar, B. (2001). Face covers and the genus problem for apex graphs. *Journal of Combinatorial Theory, Series B*, 82(1):102 – 117.

Peters, D. (2017). Recognising multidimensional Euclidean preferences. In *Proceedings of AAAI-2017*, pages 642–648.

Peters, D. and Lackner, M. (2020). Preferences single-peaked on a circle. *Journal of Artificial Intelligence Research*, 68:463–502.

Peters, D. and Skowron, P. (2020). Proportionality and the limits of welfarism. In *Proceedings of EC-2020*, pages 793–794.

Phragmén, E. (1894). Sur une méthode nouvelle pour réaliser, dans les élections, la représentation proportionnelle des partis. *Öfversigt af Kongliga Vetenskaps-Akademiens Förhandlingar*, 51(3):133–137.

Plott, C. (1967). A notion of equilibrium and its possibility under majority rule. *The American Economic Review*, 57(4):787–806.

Procaccia, A., Rosenschein, J., and Zohar, A. (2008). On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362.

Skowron, P., Faliszewski, P., and Lang, J. (2016). Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence*, 241:191–216.

Skowron, P., Faliszewski, P., and Slinko, A. (2019). Axiomatic characterization of committee scoring rules. *Journal of Economic Theory*, 180:244–273.

Thiele, T. (1895). Om flerfoldsvalg. In *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger*, pages 415–441.

Valiant, L. (1981). Universality considerations in VLSI circuits. *IEEE Transactions on Computers*, 30(2):135–140.

Appendix

1D-VCR Elections

Below we show an example of a 1D-VCR election which is neither 1D-VR nor 1D-CR.

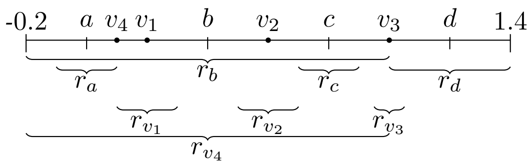
Example 1. *Let* $C = \{a, b, c, d\}$ *be the set of candidates, and* $V = \{v_1, v_2, v_3, v_4\}$ *be the set of voters. The voters and the candidates are located in the following points:*

$$\begin{array}{llll} v_1 = 0.2, & v_2 = 0.6, & v_3 = 1, & v_4 = 0.1, \\ a = 0, & b = 0.4, & c = 0.8, & d = 1.2. \end{array}$$

The radii of the candidates and voters are as follows:

$$\begin{array}{llll} r_{v_1} = 0.1, & r_{v_2} = 0.1, & r_{v_3} = 0.05, & r_{v_4} = 1.1, \\ r_a = 0.1, & r_b = 0.6, & r_c = 0.1, & r_d = 0.2. \end{array}$$

We present this election visually below:



It is easy to verify that the approval sets are:

$$\begin{array}{ll} A(v_1) = \{a, b\}, & A(v_2) = \{b, c\}, \\ A(v_3) = \{b, d\}, & A(v_4) = \{a, b, c, d\}. \end{array}$$

To see that the election is not 1D-VR, note that if it were, then the nearest candidates next to b *would have to be* a*,* c*, and* d *(because for each of these candidates there is a voter that approves him or her together with* b*). However, this is impossible, because only two candidates can be placed next to* b *(one on each side).*

To see that the election is not 1D-CR, we note that if it were, then, equivalently, the election would have to have the voter interval property. However, this would mean that there is an order such that each of v_1*,* v_2*, and* v_3 *are placed right next to* v_4*. Since only two voters can be right next to* v_4 *in any order, this is impossible.*

Total Unimodularity and 1D-Euclidean Preferences

Peters and Lackner (2020) designed an algorithm for computing winning committees according to some Thiele methods (such as PAV) in the 1D-VR model. The algorithm is based on integer linear programming formulation, admitting a relaxation which is totally unimodular if preferences are 1D-VR, and which thus admits an integral optimal solution. Our examples below demonstrate that this technique is not applicable to 1D-Euclidean preferences.

A matrix $A = (a_{ij})_{i \le n; j \le m} \in \mathbb{Z}^{n \times m}$ with all $a_{ij} \in \{-1, 0, 1\}$ is *totally unimodular* if every square submatrix B of A has $\det(B) \in \{-1, 0, 1\}$. A binary matrix $A = (a_{ij})_{i \le n; j \le m} \in \{0, 1\}^{n \times m}$ has the *strong consecutive ones property* if the 1-entries of each row form a connected block, i.e., an interval. A binary matrix has the *consecutive ones property* if its columns can be permuted so that the resulting matrix has the strong consecutive ones property. Every binary matrix with the consecutive ones property is totally unimodular. Peters and Lackner (2020) observed that for any 1D-VR election, the binary matrix of preferences has the consecutive ones property, and, thus, is totally unimodular. By some well-known facts on preservation of total unimodularity under certain matrix operations (cf. Propositions 2, 6 and 9 in Peters and Lackner (2020)) it follows that if input preferences are VR 1D-Euclidean, then the constraint matrix for Integer Programming (IP) formulation of PAV or CC is totally unimodular. However, if we consider the 1D-VCR model, without restriction to VR (or CR), then even the preference matrix might not be totally unimodular, as illustrated in the following example.

Example 2. *Consider the election from Example 1. The preference matrix of this election is (the rows correspond to the voters $v1_, \dots, v_4$ and the columns correspond to the candidates, a, b, c, d):*

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

and its determinant is -2. This violates the total unimodularity condidtion.

Furthermore, even if we consider the 1D-CR model, corresponding to the voter interval (VI) domain restriction, the method based on total unimodularity does not work either. To see this, consider the following example.

Example 3. *Let $V = \{v_1, v_2, v_3\}$, and $C = \{a, b, c, d\}$, and*

$$A(v_1) = \{a, b\}, A(v_2) = \{b, c\}, A(v_3) = \{b, d\}.$$

This election exhibits the VI property. Its preference matrix M is totally unimodular, but the matrix constructed from M by adding a row of 1s, which is the constraint matrix both for the IP-PAV and IP-CC, is identical to the preference matrix from Example 2, and as such, it is not (totally) unimodular.

Missing Proofs

In this section we present the proofs that were omitted from the main part of the paper and two corollaries to the main theorems that shed some additional light on the results.

Proposition 2. *There is an algorithm that given a 1D-VCR election and committee size, computes some winning CC committee in polynomial time.*

Proof. Consider a 1D-VCR election $E = (C, V)$, with $C = \{c_1, \dots, c_m\}$ and $V = \{v_1, \dots, v_n\}$, and committee size k ($k \leq m$). Each agant $a \in C \cup V$ has position x_a and radius r_a. We convert this representation so that for each agent a we have interval $[b(a), e(e)]$ where $b(a) = x_a - r_a$ and $e(a) = x_a + r_a$. A voter approves a candidate if their intervals have nonempty intersections. W.l.o.g., we assume that there are no two candidates so that the interval of one is fully contained in the other (indeed, every voter who would approve the candidate with the smaller interval would also approve the one with the larger). We also assume that the candidates are ordered so that $b(c_1) < \cdots < b(c_m)$ (we have strict inequalities due to the previous assumption).

For each $i \in [m]$ and $j \in [k]$, let $f(i, j)$ be the highest CC score possible to obtain by a committee of size j that consists of candidates from the set $\{c_1, \dots, c_i\}$ and includes c_i; we let

$f(i,j) = \infty$ if a committee satisfying these requirements does not exist. Note that $f(i,1)$ can be computed in polynomial time for each $i \in [m]$. For $i > 1$, we express $f(i,j)$ recursively as:

$$f(i,j) = \max_{\ell \in [i-1]} f(\ell, j-1) + |\{v \in V(c_i) \mid b(v) > e(c_\ell)\}|.$$

To understand this formula, note that, due to our assumptions, all voters v with $b(v) \leq e(c_\ell)$ either approve c_ℓ (so we should not account them due to including c_i in the committee) or do not approve c_i.

We can compute function f in polynomial time using standard dynamic programming approach. The score of the winning size-k committee for our election is $\max_{i\in[m],j\in[k]} f(i,j)$. We can compute a committee with this score using standard methods. $\square$

Next we show a somewhat stronger variant of Theorem 1, for the case of PAV.

Proposition 3. *Finding a winning committee according to PAV is* NP*-hard for 2D-VR elections, even if all voters $v \in V$ have the same approval radius r_v and the same size of their approval sets $A(v)$.*

Proof. The statement holds for any w-Thiele rule, for which $2w(1) > w(2)$ (so in particular, it holds for PAV as in its case we have $w(1) = 1$ and $w(2) = 3/2$). As in the proof Theorem 1, we reduce from the Independent Set problem for almost integral penny graphs (where the graph is given by its geometric representation). Let (G, r) be an instance of IS, where G is an almost integral penny graph and r is an integer. Let n denote the number of edges in G. We distinguish the following points in $\mathbb{R}^2$ (we will use them as the ideal points of the agents):

Direct Vertex Points: For each vertex x_i that is of degree three (i.e. which is not intermediate in G), we have a vertex point located in the center of x_i's disk; we overload the notation and also refer to this point as x_i.

Intermediate Vertex Points: We distinguish the vertex points that correspond to intermediate vertices. (i.e., those of degree two).

Edge Points: For each edge $e = \{x_i, x_j\}$ in G, we have a point in the middle of e, to which we refer as e_{ij} (we view edges in G as straight, unit-length line segments).

Intermediate Dummy Points: For each intermediate vertex x_i, denote by x_{i_1} and x_{i_2} the vertices that x_i is adjecent to. Consider the line orthogonal to the edges connecting x_i to these two vertices, i.e., take the bisector of the line segment $\overline{x_{i_1}x_{i_2}}$ and let q_i on this bisector (on an arbitrary side) such that its distance from the line segment (and, since x_i lies exactly in the middle of $\overline{x_{i_1}x_{i_2}}$, from x_i) is ϵ, for any $\epsilon \in (0, \frac{1}{2})$.

Given I, we construct an election $E = (C, V)$ with the following candidates and voters:

1. The set of candidates consists of *vertex candidates* and *dummy candidates*: In each vertex point x_i we put one vertex candidate, called c_i, and in each dummy intermediate point q_i we put one dummy candidates, called d_i.

2. We have the following two groups of voters: the *edge voters*: for each edge $e = \{x_i, x_j\}$ we have a voter located in point e_{ij}, who approves c_i, c_j, and the *dummy voters*: in each intermediate dummy point q_i we put one voter u_i, who approves the candidates c_i and d_i.

We set the committee size to be $k = r$. We ask if there exists a committee with w-score greater or equal to $s = 3r \cdot w(1)$. This ends the construction.

Let us now show that if all the voters have approval radius $1/2$ then they approve exactly the candidates indicated above. First, observe that for each non-intermediate vertex x_i of the penny graph, the distance between point x_i and each dummy intermediate point q_i (i.e., each dummy intermediate point associated with an intermediate vertex adjecent to x_i) is at least $\sqrt{1^2 + \epsilon^2} > 1/2$. Thus the distances between the non-intermediate vertex voters and the dummy candidates, as well as between the non-intermediate vertex candidates and the dummy voters, are strictly greater than $1/2$, and the respective approval ballots do not interfere with each other. Secondly, the distance between any edge voter v_{ij} and a dummy candidate c_i is at least $\sqrt{(1/2)^2 + \epsilon^2} > 1/2$. Further, it is clear that if ϵ is sufficiently small, then for each q_i, its distance from each non-intermediate vertex x_ℓ with $\ell \neq i$ is also strictly greater than $1/2$.

The distance between any intermediate vertex voter v_i and dummy candidate d_i is exctly $\epsilon < 1/2$. The same holds for the distance between dummy voter u_i and intermediate vertex candidate c_i. Obviously, dummy candidate d_i and dummy voter u_i are located at the same point. This completes the proof that the voters' approval sets are indeed as indicated in the construction, and in particular, each voter approves exactly two candidates.

It remains to show that the reduction is correct. Let us assume that G is a "yes" instance of IS. Take any independent set U of size r and define the corresponding committee to be $S_U = \{c_i \mid x_i \in U\}$, i.e., let S_U consist of the vertex candidates corresponding to the members of U. Let us calculate the w-score of S_U: each candidate from the committee contributes $3w(1)$ points. Indeed, each of them is approved by three voters. Every non-intermediate candidate is approved by three edge voters, and each intermeidate vertex candidate is approved by two edge voters and one dummy voter. So the w-score of S_U is $3r \cdot w(1)$, as required.

For the other direction, assume that there is a size-k committee S with score at least $s = 3r \cdot w(1)$. Each vertex candidate (intermediate or not) is approved by exactly three voters and, thus, can contribute at most $3w(1)$ points to the score of the committee. However, each dummy candidate d_i can only contribute at mots $w(1)$ points to the score of the committee, as he or she is approved only by the dummy voter d_i, so the committee S must consist of only vertex candidates. Further, if vertices corresponding to the ideal points of the members of S did not form an independent set, then at least one of the candidates from S would have contributed at most $w(1) + w(2) < 3w(1)$ points to the score of the committee. Thus, if S achieves a score at least $s = 3r \cdot w(1)$, then the candidates S have to form an independent set. $\square$

Theorem 2. *Deciding if there is a MAV committee of a given size and with at most a given score is* NP*-hard for 2D-VR elections (2D-CR elections) even if all voters (all candidates) have the same approval radius and each voter approves the same number of candidates.*

Proof. We reduce from the Penny Graphs Vertex Cover problem (PGVC), which is known to be NP-hard (Cerioli et al., 2011, Theorem 1.2). Let I be an instance of PGVC, where we are given a penny graph $G = (X, E)$, and an integer k. We ask if there is a vertex cover $U \subseteq X$ of G of size at most equal to k.

From I we construct an election instance as follows. We define the set of candidates to be $C = \{c_i \mid x_i \in X\}$, i.e., the centers of the unit disks representing the vertices of the graph G are the candidates. Further, for each edge $\{x, y\} \in E$ we add voter v_{xy} located in the middle of the edge (i.e., in the middle of the unit-length line segment connecting the centers of x and y), with approval radius $1/2$ (thus the voter approves candidates c_x and c_y). We ask if there is a size-k committee whose Hamming distance to each voter is at most k.

As we have a penny graph, for each voter v_{xy} and each candidate c_z with $z \notin \{x, y\}$, the distance from v_{xy} to c_z is sufficiently large so that v_{xy} does not approve c_z. Indeed, $d(v_{xy}, c_z) \geq \frac{\sqrt{3}}{2}$.

Assume that I is a "yes" instance of PGVC and let $U \subseteq X$ be a vertex cover of G. Let $S_U \subseteq C$ consist of candidates corresponding to the vertices forming the cover of G. Then, every voter v_{xy} approves of at least one member of the committee S_U. Additionally, there might be at most $k - 1$ candidates in S_U that v_{xy} does not approve of. Therefore, the Hamming distance of S_U from an arbitrary voter v_{xy} is at most:

$$d_{\mathrm{Ham}}(A(v_{xy}), S_U) \leq 1 + k - 1 = k.$$

For the other direction, suppose there is no vertex cover of G. Thus, for any set $U \subseteq X$ of size k there is an edge $\{x, y\} \in E$ such that $x, y \notin U$. Take the edge $\{x, y\}$ that is not covered by U and consider the corresponding voter v_{xy}. Since $c_x, c_y \notin S_U$, and for any $z \in U$, the corresponding candidate is not approved by v_{xy}, we have that $d_{\mathrm{Ham}}(v_{xy}, S_U) = |A(v_{xy})| + |S_U| = 2 + k$, which proves that for any committee $S \subseteq C$ of size k, the Hamming distance of this committee from some voter is strictly grater than k. This ends the proof. □

The same reduction also proves hardness for the Chamberlin-Courant rule.

Corollary 1. *Finding a winning committee according to CC is* NP*-hard in the 2D-VR elections (2D-CR elections), even if all voters $v \in V$ have the same approval radius r_v and the same size of their approval sets $A(v)$.*